Encoding natural priors in neural populations

Benjamin Poole
<ben@cmu.edu>

Advisor: Tai Sing Lee

SCS Undergraduate Thesis
Carnegie Mellon University

April 12, 2011

Abstract

Bayesian theories of the brain have provided insights into perception, but the underlying neural mechanisms which could implement these computations remains unknown. To perform Bayesian inference, sensory information must be combined with prior information about the natural world. We investigated how these natural priors could be learned and encoded in populations of neurons in primary visual cortex. We found that the distribution of neuronal tuning properties for depth-tuned neurons was very similar to the distribution of depths occurring in natural scenes. This finding is consistent with the hypothesis that neurons are performing optimal sampling of the natural environment based on the information maximization principle. By using the priors encoded in the tuning properties of neuronal populations, we were able to develop a framework for performing Bayesian inference in the brain.

Chapter 1

Introduction

1.1 Overview

Numerous psychological studies have shown that human observers are able to optimally integrate noisy sensory stimuli with prior information and resolve ambiguity [9, 6]. These studies show that humans can perform optimal Bayesian inference in many tasks, and contexts. However, the neural mechanisms which underly these computations remains unknown. To perform Bayesian inference, the brain must store and utilize priors, which encode our prior beliefs, knowledge, and experience about the world. Here we evaluate the encoding of prior information in primary visual cortex, and the role that this prior plays on information encoding.

Before we can adress how priors could be encoded, we must have a better grasp on the representations used by neural populations in sensory systems such as primary visual cortex. Sensory systems in the brain play the crucial role of connecting an individual to the outside world. These systems have the job of encoding and representing external stimuli by converting analog input (such as brightness of light on the retina) into binary spiking responses. The set of spiking responses of a population of neurons is thought of as the neural representation, or neural code for that population. Understanding what code is used by populations of neurons, and what information they encode remains an open question in neuroscience. Until we understand how the activity of neurons encodes information, we cannot reliably understand the types of computations being performed by the brain.

One of the tremendous difficulties in deciphering the neural code is the tremendous number and diversity of neurons in the human brain. The numerous biological

cell types, structures, and activity patterns create an incredibly complex system. Modern recording techniques can only capture a small fraction of this activity, but these samples of the population can still yield insights into the properties of the brain. When looking at a small (1mm cube) patch of primary visual cortex, there still exists a tremendous diversity in neural properties. Some neurons will respond more to colors, or textures, whereas others will respond to edges, curves, or binocular signals combined from both eyes. The diversity in properties makes the neural code much more complicated to decipher, but also allows for a more robust representation of the external world[1]. Understanding how this diversity comes about and how it aids in information coding has proved to be a difficult task.

Through evolution and learning, the brain has developed diverse neural populations which operate to encode information about the environment. But what principle has guided this learning? In other words, what is the information that this population is trying to encode, and how does its structure and diversity aid in achieving that goal? If we can learn what information a population of neurons is encoding, we will have a better idea of the function the population serves.

Many prior studies have postulated that the function of sensory systems is to provide a maximally informative representation of the world while limiting energy usage [9, 8, 7]. The classic study by Olshausen and Field demonstrated that receptive fields in primary visual cortex (V1) match the bases learned from independent component analysis of natural images [8]. These learned bases indicate that the receptive fields of brightness-tuned neurons in primary visual cortex are encoding a sparse representation of the natural images that maximizes the information content.

However, very little work has gone into relating how other properties of the natural world may influence the properties of the brain. Here we have attempted to understand how statistics of depth in the natural would are related to properties of neurons encoding depth in primary visual cortex. However depth is not explicitly encoded in visual cortex, but instead is derived from disparity, which measures the discrepancy between where an object is projected onto your left eye and where it is projected onto your right eye. In this thesis, we address the relationship between disparity in the natural world and the brain, and how priors for disparity could be encoded in neural populations.

We found that the distribution Fisher information in disparity-tuned neurons was nearly identical to the distribution of disparities in natural scenes. This result indicates that disparity-tuned neurons in primary visual cortex utilize a representation which maximizes the information about the stimulus in the external world. Furthermore, this distribution of Fisher information represents an encoding of prior information in the neural population. This bias in Fisher information may be the

brain's way of using this prior for Bayesian inference. Further work needs to address what types of algorithms could utilize this bias in Fisher information to aid in Bayesian decoding.

We also evaluated the role of temporal dynamics in information encoding, and found that disparity-tuned neurons increase the information encoded about the stimulus over time. This finding indicates that even in primary visual cortex, populations of neurons are integrating information about the stimulus over time. Thus low-level cortical areas are not just passively providing feed-forward input to higher layers, but instead may be playing a larger role in our active perception.

When looking at higher-order statistics, we found no correspondence between the co-occurrence statistics in natural rangemaps and correlations between neurons. Thus disparity-tuned neurons in primary visual cortex may not encode these more complex statistical structures.

1.2 Learning priors from nature

Deriving natural statistics from the environment requires a dataset which captures the natural world. Here we analyze 50 rangemaps collected using LIDAR. These images provide a representative example of the types of visual input present in the natural world. We selected 50 rangemaps that had resolutions near 22.5 pixels per degree, and masked out people, cars, and other unnatural objects. Typical images consisted of trees, shrubs, grass, and other natural objects.

To compare with neural data, we converted the depth from the rangemaps into disparity. This conversion requires knowledge of where humans fixate in a particular scene. Here we make the simplifying assumption that a human observer would randomly fixate on any particular pixel in the rangemap with equal probability. This greatly oversimplifies the human fixation distribution, but resembles the empirical fixation depth distribution [2]. For each fixation, we can convert the given depth map into a disparity map. These disparity maps are then a representative example of the types of inputs that the human visual system would experience.

Using these disparity maps, we can compute basic statistics such as the general distribution of disparities found in natural scenes (histogram), or look at the pairwise co-occurrence statistics between certain pixels to learn higher-order structures in disparity.

1.3 Information content of neural populations

Information from the left and right eyes is first combined in the primary visual cortex (V1) to compute disparity. In V1, large populations of neurons are found that respond preferentially to certain disparity stimuli [3]. We performed recordings from V1 of an awake behaving monkey while presenting dynamic random dot stereograms. Each trial contained a binocular movie which showed a disc at a certain disparity for 1 second. These stimuli are essentially constant disparity stimulus, and allow us to measure the firing rate of neurons as a function of disparity. Using a multielectrode array, we were able to simultaneously record from up to 50 disparity-tuned neurons. These simultaneous recordings allow us to analyze the temporal interactions and dynamics of a large group of neurons.

Here we were interested in estimating the information content of this disparity-tuned population of neurons. If the distribution of information contained in the neural population matches the distribution of disparities contained in natural scenes, then this population of neurons may be following the information maximization principle.

Directly computing the information contained in a population of neurons is intractable. To reduce the complexity of the problem, we first assumed that information is contained only in the firing rate of a neuron over the entire trial, and not in the temporal pattern of spiking activity within the trial. With this assumption, the mutual information between the stimulus and the neural response can be computed for small populations. However, we are working with hundreds of neurons and must approximate this information metric. Instead of directly estimating mutual information, we instead estimate Fisher information, which can be used to lower bound mutual information. Fisher information can be computed analytically from the means and standard deviations of neural activity.

However, this approach makes many assumptions about the variability of neural responses and provides extremely noisy results. Neurons with very low firing rates but steep slopes would provide peaks in the information, even though their tuning was extremely weak. These "bad" cells would tend to have predicted Fisher information orders of magnitude higher than the rest of the population, and would tend to have extremely low firing rates. In typical analyses these types of cells were thrown out, but we wanted to keep them in the analysis to estimate the information contained in the *entire* population.

To cope with this difficulty, we approximated the Fisher information through the Cramer-Rao bound. The Cramer-Rao bound is an inequality which states that the

Fisher information is greater than the inverse of the variance of any estimator. Thus if we can come up with an estimator for the stimulus (disparity), then we can lower bound the Fisher information contained in the neural population. Any estimator can provide a lower bound, but we seek to saturate the bound to get the most accurate measure of Fisher information. In hopes of achieving a more accurate approximation, we tried using a variety of estimators including: Support vector machines, logistic regression, Bayesian decoding, locally optimal linear estimators, and k-Nearest Neighbors. We found that SVMs achieved the smallest variance, and used the variance of its estimates to approximate the Fisher information of the neural population.

We found that the distribution Fisher information in disparity-tuned neurons was nearly identical to the distribution of disparities in natural scenes. This result indicates that disparity-tuned neurons in primary visual cortex utilize a representation which maximizes the information about the stimulus in the external world. Furthermore, this distribution of Fisher information represents an encoding of prior information in the neural population. This bias in Fisher information may be the brain's way of using this prior for Bayesian inference. Further work needs to address what types of algorithms could utilize this bias in Fisher information to optimally perform Bayesian inference.

1.4   Neural dynamics and correlations

One major flaw in the previous analysis is that spike count data is aggregated over the course of an entire second. Our perceptual capabilities operate at a much faster speed, and thus must be supported by a neural code which can perform inference at short timescales [4]. In chapter 3, we analyzed the information content of our neural population as a function of time. Instead of using spike counts from entire trials, we analyzed 5 to 50ms bins of neural activity over the full 1000ms trial. To model this time-varying neural response we used Generalized linear models (GLMs), which can be used to represent the response of an individual neuron as a linear combination of other factors. We used this model to predict the spiking activity of a neuron over time using the local field potential as well as the spiking history of other neurons in the population [5]. This model was able to capture a large deal of the variability in a neuron's spiking response, and provided more accurate estimates of the information content of a neural population over time. We also used the spike-count based techniques by sliding a window over time and normalizing the firing rate. However, these techniques performed poorly in the temporal domain due to the complex dynamics of neurons. In particular, these

models fail to capture the refractory period which occurs after a neuron spikes, and inhibits it from firing again immediately.

In chapter 4, we extend the work on relating natural scene statistics to neural properties by looking at the second-order statistics of pixel co-occurrence and pairwise correlations. Based on our findings in chapter 2, we anticipated that neural populations may be optimally encoding higher-order structures such as surfaces and planes in the 3D world. For example, if surfaces are generally tilted away from the viewer (i.e the ground plane), then neurons which encode the top and bottom of the visual field may be negatively correlated. To evaluate this hypothesis, we first computed co-occurrence statistics from the rangemap data. This allowed us to create a prior over pairs of disparities, instead of just a single disparity as in chapter 2. To compare this prior with the neural data, we must estimate the connectivity between each pair of neurons. An estimate of the connectivity, or coupling strength, between pairs of neurons can be found using the GLM used in chapter 3. The coupling strength provides a measurement of the influence of one neuron on another, but we cannot know whether these neurons are anatomically connected. We found that there was no correspondence between the predicted correlation derived from the co-occurrence statistics, and the measured coupling strength determined by the GLM. This finding indicates that disparity-tuned neurons in primary visual cortex may not encode these statistical structures. Instead, it may be the job of neurons in higher-level cortical areas to capture these more complex relationships and encode the associated priors.

Bibliography

[1] Bruno B Averbeck, Peter E Latham, and Alexandre Pouget. Neural correlations, population coding and computation. *Nature reviews. Neuroscience*, 7(5):358–66, May 2006.

[2] Alan C Bovik and Lawrence K Cormack. Disparity statistics in natural scenes. *Journal of Vision*, 8:1–14, 2008.

[3] B G Cumming. An unexpected specialization for horizontal disparity in primate primary visual cortex. *Nature*, 418(6898):633–6, August 2002.

[4] Adam L Jacobs, Gene Fridman, Robert M Douglas, Nazia M Alam, Peter E Latham, Glen T Prusky, and Sheila Nirenberg. Ruling out and ruling in neural codes. *Proceedings of the National Academy of Sciences of the United States of America*, 106(14):5936–41, April 2009.

[5] Ryan C Kelly, Matthew a Smith, Robert E Kass, and Tai Sing Lee. Local field potentials indicate network state and account for neuronal response variability. *Journal of computational neuroscience*, January 2010.

[6] David C Knill and Alexandre Pouget. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends in neurosciences*, 27(12):712–9, 2004.

[7] M S Lewicki and T J Sejnowski. Learning overcomplete representations. *Neural computation*, 12(2):337–65, February 2000.

[8] Bruno a Olshausen and David J Field. Sparse coding of sensory inputs. *Current opinion in neurobiology*, 14(4):481–7, August 2004.

[9] Eero P Simoncelli. Optimal Estimation in Sensory Systems. *New York*, 2009.